NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

EVALUATING TROPICAL CYCLONE FORECAST TRACK UNCERTAINTY USING A GRAND ENSEMBLE OF ENSEMBLE PREDICTION SYSTEMS

by

Douglas W. Pearman

September 2011

Thesis Co-Advisors: Patrick A. Harr
Eva D. Regnier

Approved for public release; distribution is unlimited

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2011	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE
Evaluating Tropical Cyclone Forecast Track Uncertainty Using A Grand Ensemble Of Ensemble Prediction Systems

5. FUNDING NUMBERS

6. AUTHOR(S) Douglas W. Pearman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ______N/A__________.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)
The skill of a combined grand ensemble (GE), which is constructed from three operational global ensemble prediction systems (EPS), is evaluated with respect to the probability forecast of a tropical cyclone (TC) being within a specified area. Anisotropic probability ellipses are defined from the GE to contain 68% of the ensemble members. Forecast reliability is based on whether the forecast verifying position is within the ellipse. A sharpness parameter is based on the size of the GE-based probability ellipse relative to other operational forecast probability ellipses. For the 2010 Atlantic TC season, results indicate that the GE ellipses exhibit a high degree of reliability whereas the operational probability circle tends to be over-dispersive. Additionally, the GE ellipse tends to be sharper than the operational product for forecast intervals beyond 48 hours. The size and shape of the GE ellipses varied with TC track types, which suggests that information about the physics of the flow-dependent system is retained whereas isotropic probability ellipses may not reflect variability associated with track type. It is concluded that the GE probability ellipse demonstrates utility for combined EPS to enhance probabilistic forecasts for use as TC-related decision aids, as there is a potential for reducing the sizes of warning areas.

14. SUBJECT TERMS THORPEX, TIGGE, GPCE, Tropical Cyclone, Spatial variability, Grand Ensemble, Probability Ellipse, Flow-dependent system.

15. NUMBER OF PAGES
63

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

EVALUATING TROPICAL CYCLONE FORECAST TRACK UNCERTAINTY USING A GRAND ENSEMBLE OF ENSEMBLE PREDICTION SYSTEMS

Douglas W. Pearman
Lieutenant, United States Navy
B.A., Thomas Edison State University, 2006
M.S., University of Southern Mississippi, 2008

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY AND OCEANOGRAPHY

from the

NAVAL POSTGRADUATE SCHOOL
September 2011

Author: Douglas W. Pearman

Approved by: Patrick A. Harr
Thesis Co-Advisor

Eva Regnier
Thesis Co-Advisor

Wendell A. Nuss
Chair, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The skill of a combined grand ensemble (GE), which is constructed from three operational global ensemble prediction systems (EPS), is evaluated with respect to the probability forecast of a tropical cyclone (TC) being within a specified area. Anisotropic probability ellipses are defined from the GE to contain 68% of the ensemble members. Forecast reliability is based on whether the forecast verifying position is within the ellipse. A sharpness parameter is based on the size of the GE-based probability ellipse relative to other operational forecast probability ellipses. For the 2010 Atlantic TC season, results indicate that the GE ellipses exhibit a high degree of reliability whereas the operational probability circle tends to be over-dispersive. Additionally, the GE ellipse tends to be sharper than the operational product for forecast intervals beyond 48 hours. The size and shape of the GE ellipses varied with TC track types, which suggests that information about the physics of the flow-dependent system is retained whereas isotropic probability ellipses may not reflect variability associated with track type. It is concluded that the GE probability ellipse demonstrates utility for combined EPS to enhance probabilistic forecasts for use as TC-related decision aids, as there is a potential for reducing the sizes of warning areas.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ATCF	Automated Tropical Cyclone Forecast system
ATE	Along-Track Error
COR	Conditions of Readiness
CXML	Cyclone XML
DET	Deterministic
ECMWF	European Center for Medium-Range Weather Forecasts
EMN	Ensemble Mean
EPS	Ensemble Prediction System
ETBV	Ensemble Transform Bred Vector
ETKF	Ensemble Transform Kalman Filter
FTE	Forecast-Track Error
GE	Grand Ensemble
GFDL	Geophysical Fluid Dynamics Library
GFS	Global Forecasting System
GPCE	Goerss Prediction Consensus Error
GPCE-X	Goerss Prediction Consensus Along/Across-track error
JTWC	Joint Typhoon Warning Center
MOGREPS	Met Office Global and Regional Ensemble Prediction System
NCEP	National Center for Environmental Prediction
NHC	National Hurricane Center
OFCL	Official Forecast
SV	Singular Vectors
TC	Tropical Cyclone
TD	Tropical Depression
TS	Tropical Storm
THORPEX	The Observing System Research and Predictability Experiment
TIGGE	THORPEX Interactive Grand Global Ensemble
UKMET	United Kingdom Meteorological office
WMO	World Meteorological Organization
WWRP	World Weather Research Program
XTE	Cross-Track Error

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

Thank you to my wife, Holly, my daughters, Cienna and Hannah, and the rest of my family who have encouraged me and showed patience through missed activities during this academic effort.

I would also like to thank my advisors, Professor Pat Harr and Professor Eva Regnier for accepting me as a thesis student at a time when my health was not certain and all I brought was a somewhat interesting idea. Your encouragement and interest as we put meat on the bones of this project helped get through times when things were murky.

Thanks also to Captain Stone for taking time away from your responsibilities to have countless discussions about statistics and physics of flow-dependent systems. Your willingness to assist any student who comes to your door is exemplary.

To my friends, thanks for the constant and sometimes not so gentle reminders that there is life outside of work. Your innovative ideas on how to maximize the Monterey experience were much appreciated!

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. MOTIVATION

Tropical Cyclones (TCs) pose a seasonal threat to oceanic transit lanes and to coastal military installations located in the climatological paths of TCs (USN 2010a). The sortie of mobile assets and preparing shore-based installations for damaging wind and floodwaters can reduce the impact of TCs on the United States Navy fleet readiness. These expensive risk management decisions are made based on guidance from operational TC forecasting centers.

The impact on lives, readiness, and the cost of sending the fleet to sea is staggering. Sortie of the ships in Norfolk, VA in advance of Hurricane Floyd in 1999 is estimated at $36M (*Norfolk Virginian-Pilot*, 27 September 2003). A Fleet Commander follows a risk-based decision process to set hurricane evacuation conditions of readiness (COR) that are based on TC forecast tracks and radii of destructive wind. For purposes of this thesis, it is assumed that sustained winds of 50 kt is considered destructive and will serve as the risk threshold for sortie of the fleet. The cost of an unnecessary sortie is high, but leaving ships moored through passage of a TC could be devastating.

The decision to sortie the fleet is recommended by a Fleet Oceanographer using personal forecasting skills guided by operational forecast products from the National Hurricane Center (NHC), Joint Typhoon Warning Center (JTWC), and other basin-specific national forecast centers. The Fleet Oceanographer assesses each sequential forecast scenario and makes recommendations based on scientific assessment of the information available at a pre-defined threshold cutoff time in advance of the approaching TC. Among various forecasting techniques, existing guidance from the TC forecast centers might be based on a consensus of operational numerical deterministic forecast models. To convey variability among models, the forecast center adds a seasonally static isotropic error swath around the consensus mean. The consensus mean may be defined using a variety of models and competing factors. A similar display encompasses the official forecast track that is constructed by the operational TC forecaster after appropriate examination of all available aids, which are based on numerical, statistical, and observational information. A more detailed account of TC track and product generation will be discussed in Chapter II.

In this thesis, an assessment of forecast uncertainty is examined. The goal is then to improve on the characterization of uncertainty by communicating confidence in individual forecasts using dynamic probability ellipses. These probability ellipses are based on a Grand Ensemble (GE) of operational ensemble prediction systems (EPS) as a measure that represents the limit of predictability of a given TC forecast scenario.

In 2005 the World Meteorological Organization/World Weather Research Program (WMO/WWRP) implemented The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) database. The TIGGE database is populated with ensemble products from ten national operational weather centers (Bougeault et al. 2011). These data are available for operational use as well as delayed availability for research. Dynamic probability guidance presented by a GE presents an efficient addition to the existing tools available to meteorological professionals making risk management recommendations associated with TC forecast tracks.

In early September, 2010, Hurricane Earl approached the eastern seaboard of North America. The forecast track indicated a northeastward turn slightly seaward of Norfolk, VA. The operational weather center in Norfolk and Commander Second Fleet Oceanographer’s staff presented commanders with probability swaths produced by NHC. These swaths were generated based on 1000 statistically plausible scenarios derived from the probability distribution of past error (DeMaria et al. 2009 and Split et al. 2010). With each successive forecast, the swaths indicated Norfolk to be in an area of declining probability of experiencing destructive winds. This supported the forecaster’s recommendation and directly led to the non-sortie decision (CDR S. Memmen, C2F Oceanographer, 2009, personal communication). Hurricane Earl remained offshore and the application of the stochastic decision aid was considered a success. This scenario demonstrates how conveyance of confidence in a TC forecast presented in a probabilistic framework can guide destructive weather risk management decisions. However, the basis for probabilistic determination of forecast uncertainty may be rooted in a number of constructs. DeMaria et al. (2009) defines the process in use at NHC based on a stochastic sampling from predefined distribution of historical forecast errors. However, recent results from Majumbdar and Finochio (2010) suggest that a collection of dynamically-generated ensemble members may span the set of possible outcomes equally well. Furthermore, the dynamically-generated ensemble members may have other applicable qualities for TC forecasts. This research examines forecast tools based on ensemble forecast systems and the decisions they guide. In particular the assessment of reliability, sharpness and utility of a GE from EPS in the TIGGE database will be defined.

B. OBJECTIVE

Ensemble modeling and various stochastic forecasting techniques have rapidly developed into mainstream operational forecasting over the past decade (Bougault et al. 2011). The development of the TIGGE database and subsequent population of TIGGE with current EPS leads naturally to developing a dynamic tool to present confidence in a current TC forecast scenario based on statistics and physics of the existing and predicted state of the atmosphere. The GE concept capitalizes on these advances in EPS and the development of the TIGGE database. Thus, forecasters are provided with dynamic-statistical representations of the limit of predictability of a particular TC forecast track based on current physics to complement those of combined model output and historical error. The spatial distribution of the GE forecast TC positions could be depicted with an anisotropic ellipse using any percent confidence level. For purposes of this research the probability ellipse will be defined as the 68% confidence level following existing operational probabilistic products.

The objective of this thesis is to improve upon existing TC forecasting tools used to quantify uncertainty and communicate confidence in individual forecast scenarios. The following hypothesis will be investigated:

- An anisotropic error ellipse developed from a grand ensemble (GE) of three global EPS serves as a predictor of forecast uncertainty and demonstrates skill when compared methods used in current operational descriptions of forecast uncertainty.

Current probability graphics provided to Fleet Oceanographers describe the static isotropic historical error of the forecast center's predictions. These products convey little about the confidence in the forecast given the predictability of the existing atmosphere. Proposed in this thesis is a method of communicating to Fleet Oceanographers the predictability of the atmosphere based on the variable output of multiple EPS. This GE provides a single look at the confidence in a given forecast scenario based on robust combination of physics and statistics. The physics are inherent in the operational EPS and the statistics are calculated from the output of the stochastic predictions. This study provides insight into the utility of a GE in communicating confidence in a forecast track and assesses how this information will assist in the sortie decision process when compared to existing operational probability graphics.

Background material is provided in Chapter II. The analysis methodology is described in Chapter III. The results are presented in Chapter IV. A case study and conclusions with future recommendations are outlined in Chapter V.

II. BACKGROUND

A. OPERATIONAL METHODS USED TO DEFINE TC FORECAST UNCERTAINTY

Many probabilistic forecast aids are available to the operational TC forecast process. While all attempt to convey the uncertainty in the atmosphere as a flow-dependent system, each has its own strengths and weakness based on the slightly different methods of developing the forecast tool. The following three sections will provide an introduction to the probabilistic tools relevant to this thesis.

1. Standard Forecast Track with Cone of Error

Tropical cyclone forecast track warnings are produced at NHC every six hours. The forecast track product consists of the TC position every 12 hours out to 120 hours, TC category (tropical depression (TD), tropical storm (TS), and hurricane (HU)), and an uncertainty cone. The uncertainty cone (Figure 1) is an area defined at each forecast interval based an isotropic representation of on two-thirds mean forecast track error and is expected to contain the forecast track 70% of the time (NHC 2010b). While the uncertainty cone serves as a viable method of communicating confidence in the NHC forecast, it does not adequately convey confidence in any particular TC forecast track. In a flow-dependent system, confidence in the predictability of the current atmosphere and TC are dependent on location in basin, track type (i.e. recurving or straight), intensity, and many other parameters. The dependencies on these factors are not represented in the definition of the NHC cone of uncertainty shown in Figure 1.

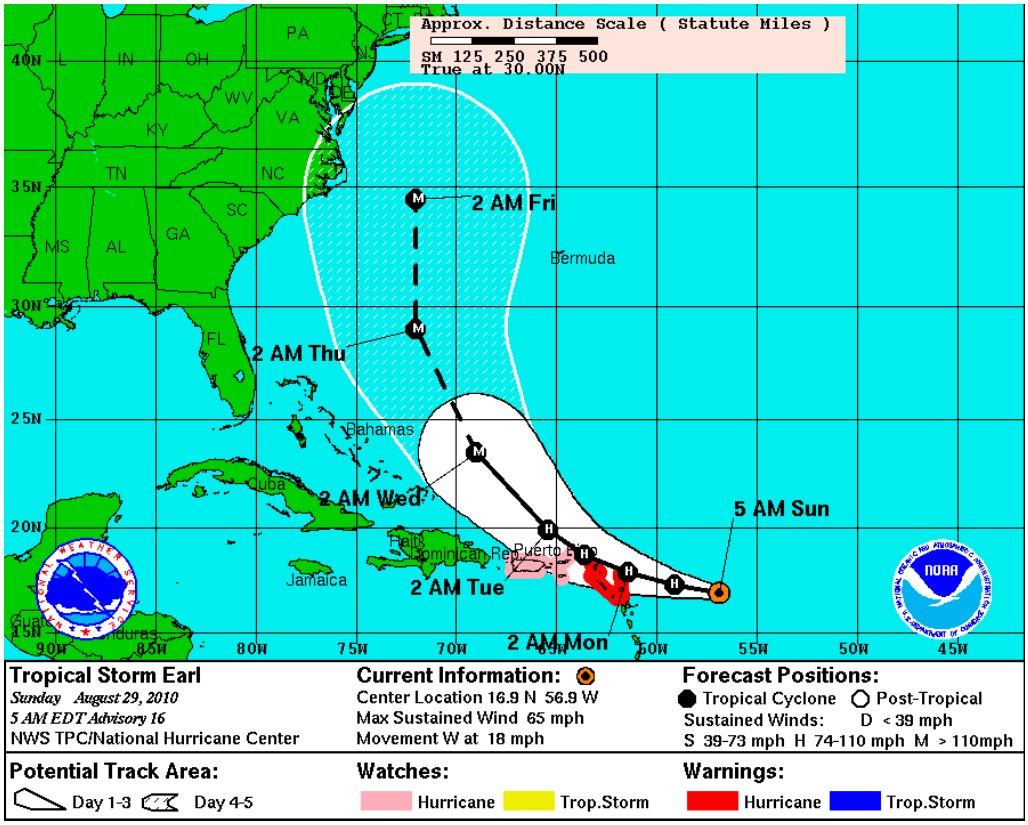

Figure 1. The NHC Cone of Uncertainty about Hurricane Earl for the forecast issued at 1500 UTC 29 August 2010. The white swath is anisotropic representation of the historical forecast uncertainty at NHC to 67% confidence.

2. Goerss Predicted Consensus Error (GPCE)

Goerss (2000) examines the average of multiple deterministic models to aid forecast centers in TC forecast track prediction. The models selected are available in time to be considered as aids to forecasters before release of the official forecast. The distribution of the ensemble of deterministic models represents the probable error of the TC forecast track at specific times. Since each deterministic model in the ensemble is generated from a different forecast center, it contains the observational error unique to that model as well as the errors in the physics from the model itself. On average, the mean of this consensus performs better than any individual model and the spread of the ensemble provides insight into the predicted forecast error (Goerss 2000).

Operational TC forecasters often use the consensus aids (Goerss 2007) as the first guess approximation for the TC forecast location. Carr and Elsberry (2000) and Sampson et al. (2006) further add to the utility of the consensus aid by suggesting a selective consensus in which individual models from the ensemble that are obvious outliers are omitted. The utility of the consensus and the error of the spread are formalized with the development of the Goerss Predicted Consensus Error (GPCE) (Goerss 2007).

The operational GPCE consists of an isotropic representation of 68% of the predicted forecast error contained within a circle calculated using linear regression techniques trained on the distribution of errors from previous years (Figure 2). The GPCE is thought to contain the most probable TC forecast location and is used operationally to form the best probable location of the predicted TC position (Goerss 2007). Although the GPCE circle is an isotropic representation of error, the distribution of errors is often anisotropic in an along and across-track sense. Thus, GPCE may not effectively communicate the orientation of the error in cases where the forecast is more confident in either along or across-track direction. It may in fact over predict error in cases where errors are distributed such that along-track or across-track errors dominate.

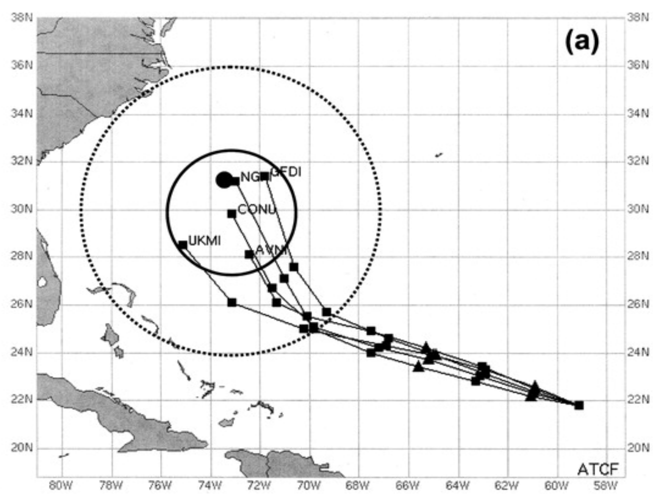

Figure 2. The predicted 68% confidence radius GPCE circle (solid circle) surrounding the 120-h consensus forecast for Hurricane Isabel, initiated at 0000 UTC 13 Sep 2003 (Goerss 2007). The individual model tracks used to create the consensus forecast track are shown along with the 120-h error radius used by NHC.

3. Goerss Predicted Consensus Error-Along/Across-track (GPCE-X)

Hansen et al. (2010) extends the GPCE product to an anisotropic ellipse with its axes defined by along and across-track errors (GPCE-X). Generally, the along-track errors (ATE) define the portion of the forecast error that is either ahead or behind the verifying position. While the across-track errors define the portion of the predicted error of the consensus that is to the left or right of the verifying position (Figure 3). In this setting the along-track (across-track) errors define the portion of the predicted error of the consensus that is ahead and behind (left and right of) the consensus mean position. This work improves upon the sharpness of the predicted area by considering along and across-track spread thereby reducing the area of the ellipse. The GPCE-X represents 70% of the predicted positions of the members contained in the consensus of deterministic models, and predicts forecast track error in an anisotropic manner in an along-track and across-track sense.

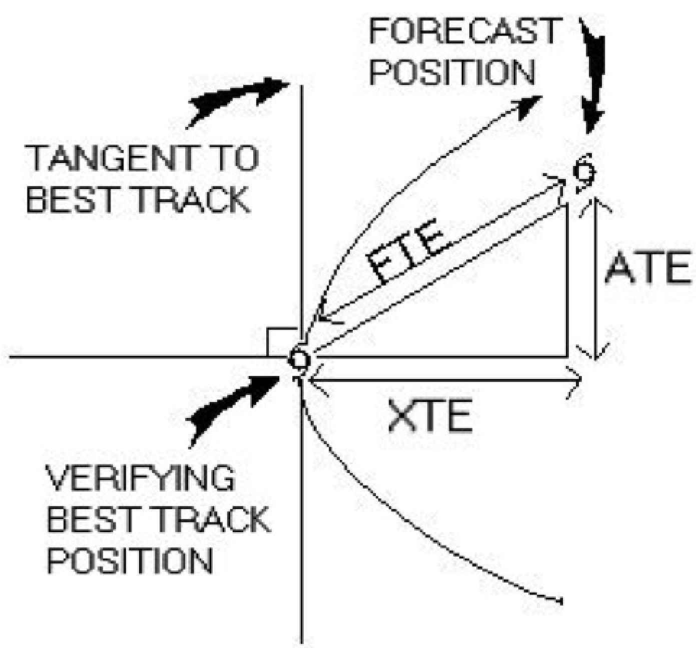

Figure 3. Definition of across-track error (XTE), along-track error (ATE), and forecast track error (FTE).

B. OVERVIEW OF SELECTED ENSEMBLE MODELS

Many operational EPS exist today. This thesis focuses on the European Center for Medium Range Weather Prediction (ECMWF), the United Kingdom Meteorological Office (UKMO) Met Office Global and Regional Ensemble Prediction System (MOGREPS), and the National Center for Environmental Prediction (NCEP) Global Forecast System (GFS). These three are the most commonly used EPS at the official forecast centers. Each of these EPS is based on slightly different approaches to defining perturbations required to generate the ensemble members. In general, the purpose of ensemble model forecasting is to define a flow-dependent representation of model forecast uncertainty that is based on the overall chaotic nature of the atmosphere and its representation by the governing equations of motion (Lorenz 1963). The following three sections provide a brief overview of how these EPS are developed.

1. European Center for Medium Range Weather Prediction (ECMWF)

The ECMWF EPS has a horizontal resolution of 32 km. One control forecast based on the operational analysis used in the deterministic model serves as control, and 50 additional members are developed from perturbed initial conditions. The EPS represents the initial uncertainty through 25 pairs of perturbations subtracted from and added to the analysis field. Perturbations are based on singular vectors (SV) that define regions of expected error growth throughout the integration of the forecast model (ECMWF 2011). The SV are selected based on the greatest linear growth rate in total energy over a 48- hour time period for a fixed set of norms, assumptions, and spatial targets (ECMWF 2011). The pertinent analysis fields are then integrated forward defining 50 differing end states.

In January of 2002, the ECMWF began including tropical singular vectors to account for the specific physical processes involved in tropical convection and the development of a tropical cyclone. These are perturbations defined using a linear physics package developed to describe model parameterizations such as sub-grid scale deep cumulous convection that dominate in the tropical environment. The resulting SV are used to define analysis perturbations in the defined TC development regions

(ECMWF 2011). In addition to the SV, the ECMWF EPS introduces model perturbations to consider effects of model error. This essentially serves as a method to simulate random uncertainty in model errors arising from the parameterizations of physical process (ECMWF 2011).

2. United Kingdom Meteorological Office (UKMO)

The UKMO EPS is called the Met Office Global and Regional Ensemble Prediction System (MOGREPS). The MOGREPS has a horizontal resolution of 60 km and consists of 24 ensemble members developed from perturbed initial conditions.

The perturbations of the model physics are similar to those in the ECMWF EPS. However, the analysis perturbations are generated using an Ensemble Transform Kalman Filter (ETKF). The ETKF method allows for rapid calculation of analysis perturbations as the computation relies on analysis perturbations being in the subspace of the forecast perturbations (Bowler et al. 2008). The error is bred for each member allowing for perturbations to be rescaled for each member separately. The analysis is sampled in such a way as to determine the observations that most control the forecast. These portions of the analysis are then perturbed. In a simple model of the ETKF, the control forecast is subtracted from the perturbed forecast and multiplied by a scaling factor, then added to the control analysis resulting in the perturbed analysis that is used for the initial conditions of each individual member (Bowler et al. 2008).

3. National Center for Environmental Prediction (NCEP)

The NCEP Global Forecasting System (GFS) uses an Ensemble Transform Bred Vector (ETBV) to perturb initial conditions. The ETBV perturbations are selected via the following steps:

- 20 perturbed initial conditions are integrated for six forecast hours.
- The 20 forecasts are differed from the analysis valid at the same time.
- The differences are made to be statistically independent and rescaled to ensure they are representative of known observational and background errors.

- This cycle repeats every six hours breeding a new set of perturbations. This process is repeated until the averaged differences become stable.

Additional perturbations known as stochastic perturbations are introduced to account for the effect of uncertainty in the numerical model on forecast variables. This is accomplished by perturbing the forecast variable total time tendency of all physical and dynamical processes by a random factor, and scaling to ensure the perturbed variable is appropriately sized for the region (UCAR 2010).

4. An Isotropic Joint-Ensemble

Majumdar and Finochio (2010) develop a probability circle from a joint ensemble of the ECMWF and the UKMO MOGREPS global EPS ensemble members. They compare this joint ensemble circle to the GPCE circle. The FTE of the joint ensemble mean is found to be comparable to the consensus mean FTE. The distribution of locations contained in the joint ensemble prove to perform comparably to GPCE. However, the distribution of forecasts contained in the joint ensemble is found to have a larger forecast area than the GPCE circle, which indicates a less sharp forecast. Majumdar and Finochio (2010) observe that the global ensemble prediction system is appealing in that it provides a predicted probability distribution. However, if it is not conditioned on previous seasons, the size of the ellipse may vary from forecast to forecast. The GPCE and GPCE-X approaches are dependent on the past several seasons. Therefore, the size is consistent from forecast to forecast, and the predicted error consistently grows with each subsequent forecast interval.

Borrowing from the GPCE model, this thesis develops the GE probability ellipse from independent global EPS. Each EPS contains independent errors unique to that EPS. The variability in the GE probability ellipse is independent of past error, which is considered a benefit that will be discussed in Chapter IV. The GE probability ellipse builds on findings of Majumdar and Finochio (2010) and Hansen et al. (2010) and results in an anisotropic depiction oriented by the spatial distribution of ensemble member positions.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODOLOGY

A. DATA

1. Data Source

The data for this study consist of forecast positions of all Atlantic TCs that achieved tropical storm strength or greater during years 2008-2010 (Figure 3, 4, 5). The dataset is limited in size because of the UKMO MOGREPS is only available for recent years. All three available years are used to define the general characteristics of the EPS forecasts (i.e. forecast track errors of the mean and hit rates), however, the availability of the GPCE and GPCE-X data restricts comparison with these products to be limited to only 2010.

The 2008 season (Figure 3) contains the most storms of the three years. The 2008 season has 16 named storms, eight of which reach hurricane strength. The 2009 TC season (Figure 4) contains the fewest named storms that reach TC strength. The season consists of nine named TCs, four of which reach hurricane strength. The 2010 Atlantic tropical cyclone season (Figure 5) is above average in activity with 19 named storms. Twelve of these TCs reach hurricane strength. In all, the dataset is made up of 44 named TCs that consists of 23 hurricanes, 16 of which are major hurricanes, which is defined by wind speeds greater than 120 kt. Eight of the TCs make landfall on the continental United States.

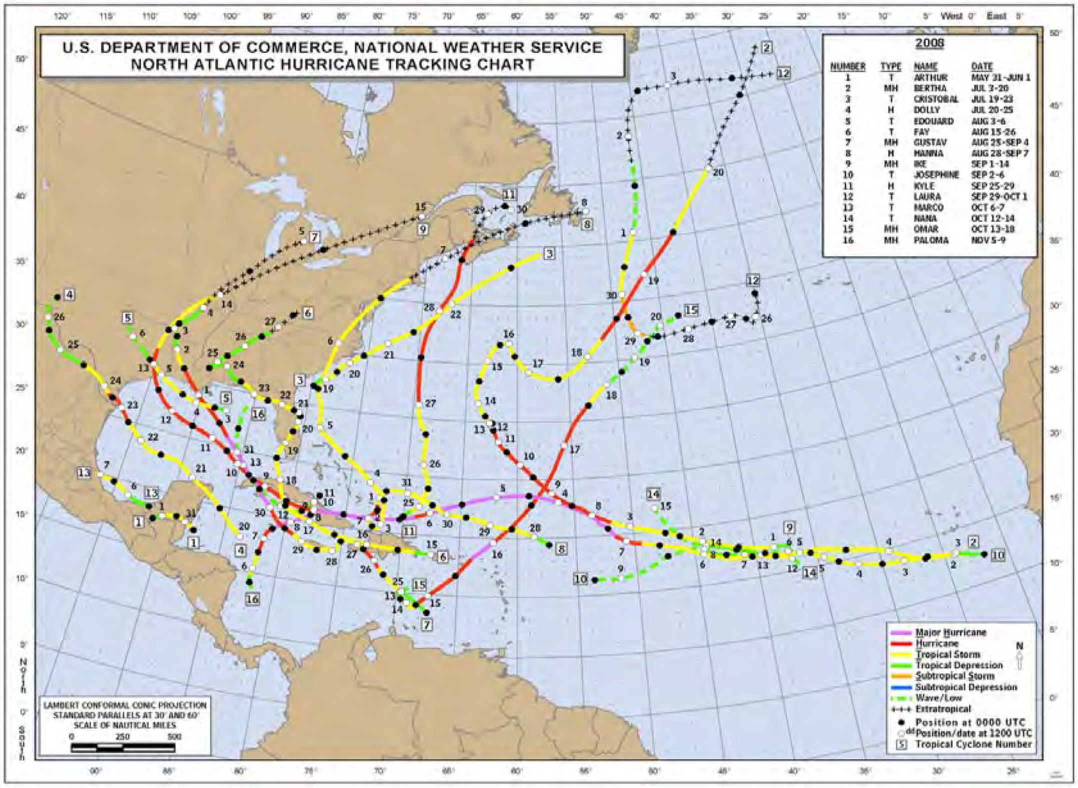

Figure 4. The 2008 Atlantic Tropical Cyclone season summary. The track information is outlined in the lower right hand corner of the graphic. (From NHC 2010)

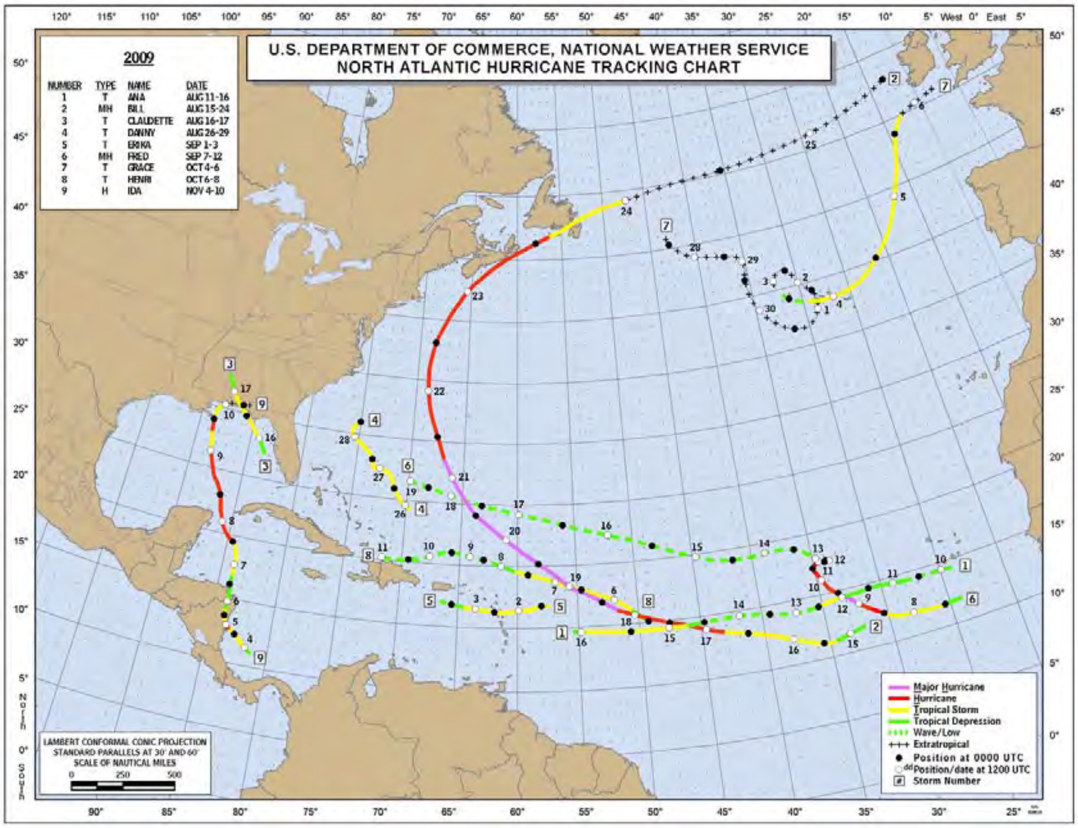

Figure 5. The 2009 Atlantic Tropical Cyclone season summary. The track information is outlined in the lower right hand corner of the graphic. (From NHC 2010)

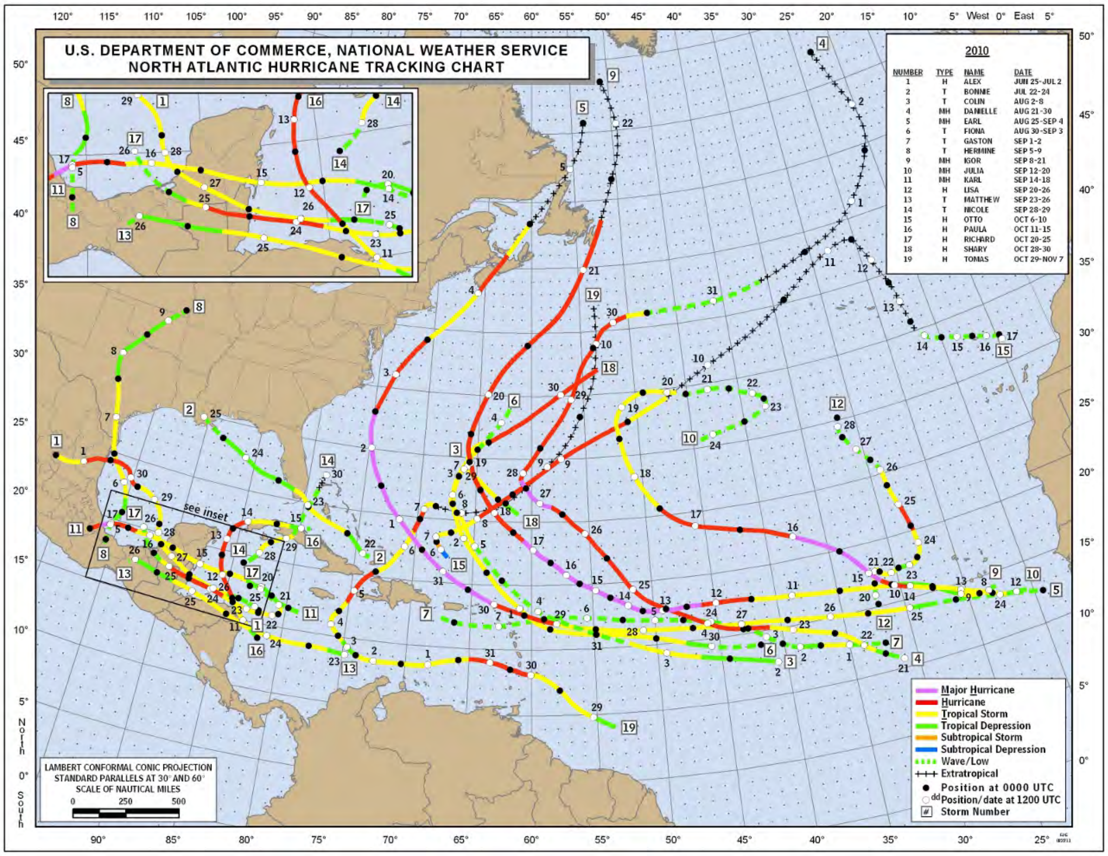

Figure 6. The 2010 Atlantic Tropical Cyclone season summary. The track information is outlined in the lower right hand corner of the graphic. (From NHC 2010)

2. Data Format

The MOGREPS and ECMWF global EPS are selected from the TIGGE database in Cyclone XML (CXML) format, which is a descriptive, human legible, standard format that is used to define forecast TC positions for all global EPS. The NCEP data are obtained from the automated TC forecast system (ATCF) and are converted to the same format of the TIGGE data to ensure standardized data processing.

The best track data are the final NHC analyzed positions for each TC used to verify forecast at each forecast time. The best track data are retrieved from the ATCF database, converted to a standard format and are used for verification of the GE at 12-h forecast intervals out to 120 h. The GPCE and GPCE-X forecast center positions and length of the circle radius are available via the Automated TC Forecast (ATCF) database. These are also converted into the standard format for ease of processing.

3. Homogeneity

Consistency through the dataset is critical for analysis. If part of a forecast is missing from one of the three EPS, the forecast for that valid time is discarded. Further, for direct comparison with the GPCE product, homogeneity is verified between the GE, GPCE and GPCE-X forecasts. What remains is a homogeneous dataset of ECMWF, MOGREPS, and GFS forecasts for each TC forecast considered. Short-lived TCs in which forecast times are not greater than 48 hours are discarded.

B. DEVELOPING THE GRAND ENSEMBLE

As defined above, forecast uncertainty may be analyzed in a variety of ways. The GPCE product characterizes uncertainty with a single measure of dispersion that is isotropic. The GPCE-X defines uncertainty using an ellipse that depends on the spread based on across-track and along-track error. In this study, the uncertainty in the GE forecast position is defined in relation to the principal axis associated with the ensemble member spatial distribution and centered relative to the GE mean position. An ellipse is defined to contain 68% of the GE ensemble member forecast track positions. The ellipse is defined to be centered on the GE mean and is generated for all homogeneous intervals of each forecast period in which there is a best track verification point and corresponding GPCE forecast.

The ellipse is defined by first creating a 2 x n matrix of the predicted latitude and longitude of each ensemble member making up the GE at a specific interval of a forecast valid time. When all members are present n=95. A covariance matrix is defined from the latitude/longitude matrix defined by:

$$\sum_{\hat{x}} = \sigma_0^2 (A^T A)^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \tag{1}$$

where A is the latitude and longitude matrix. The diagonal of the resulting covariance matrix represents the variance of each variable and the off diagonals are the covariance the variables.

As defined in Figure 7, the latitude and longitude representation of the ensemble members are rotated to a coordinate system (y_1, y_2) and covered by the ellipse as defined by the expression:

$$\frac{y_1^2}{(c\sqrt{\lambda_1})^2} + \frac{y_2^2}{(c\sqrt{\lambda_2})^2} = 1 \tag{2}$$

In Equation (2), a scale factor is represented by c, and λ_1, λ_2 are the eigenvalues of the covariance matrix. The eigenvectors of the covariance matrix serve as a rotation matrix orienting the semi-major axis of the ellipse in the direction of the largest variance, and the eigenvalues in a transformation matrix scale the semi-major and semi-minor axes. The orientation of the ellipse ϕ is defined by:

$$\tan 2\phi = \frac{2\sigma_{1,2}}{\sigma_1^2 - \sigma_2^2} \tag{3}$$

Procedurally Equations (2) and (3) are accomplished by predefining a set of generic points (P) surrounding a circle such that

$$P = 0 : \frac{\pi}{100} : 2\pi \tag{4}$$

These generic points are then transformed into the ellipse from Equation (2) by

$$[x, y] = [\cos(P^T), \sin(P^T)] * \sqrt{eigval} * eigvec^T \tag{5}$$

where eigval and eigvec are the matrices of eigenvalues and eigenvectors calculated from the covariance matrix $\sum_{\hat{x}}$. The matrix [x,y] represents a longitude and latitude matrix defining an ellipse that is centered on the origin.

Finally, the mean of the original latitude and longitude matrix is calculated to provide a center position for the ellipse. Assuming a bivariate normal distribution of the ensemble member positions, the sum of the squares represented in Equation (2) is represented by a Chi-squared distribution. A Chi-squared scaling parameter c is applied to ensure the GE ellipse captures 68% of the ensemble member forecast positions.

The following vectors then define the latitude and longitude positions of the final ellipse:

$$\begin{aligned} lon &= \mu_1 + c * x \\ lat &= \mu_2 + c * y \end{aligned} \tag{6}$$

The resulting ellipse (Figure 7) is then tested to ensure it is well calibrated in that 68% of the ensemble members were captured at 95% confidence interval over all homogenous TC dataset. As illustrated in Figure 7 the ellipse orientation retains the spatial distribution of the ensemble member positions. The semi-major axis is neither normal nor parallel to the trajectory; rather it is oriented to the spatial distribution of the ensemble members.

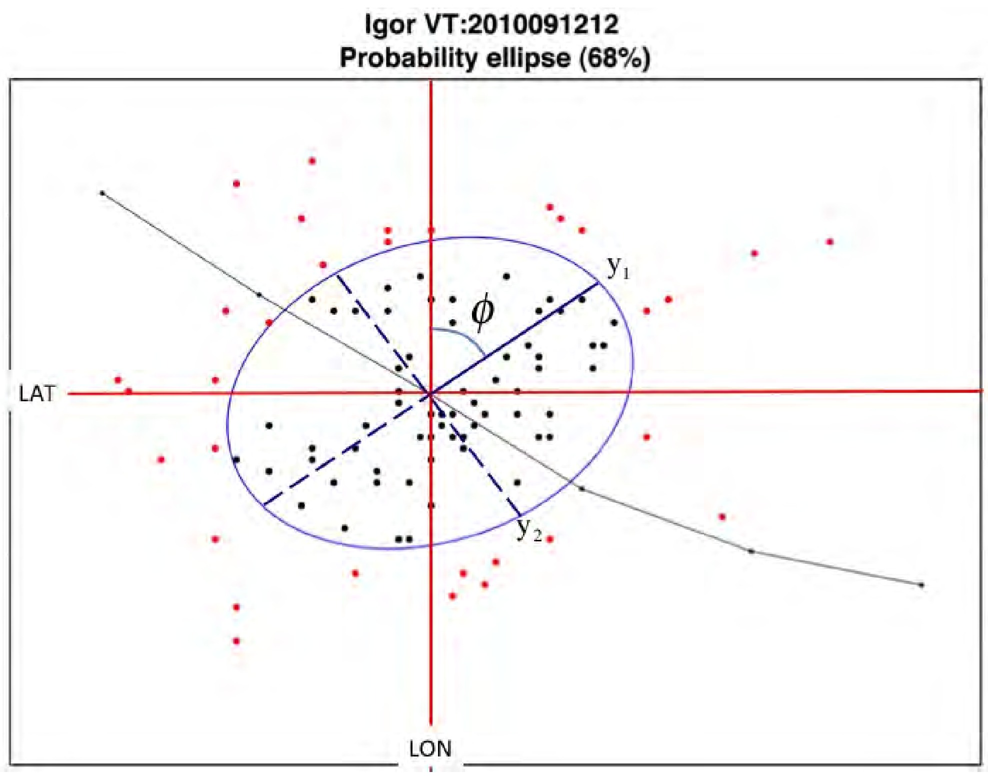

Figure 7. The probability ellipse defined for ensemble members of the 1200 UTC 12 September 2010 forecast for Hurricane Igor. The GE probability ellipse in blue contains 68% of the ensemble members (black dots) and is centered on the GE mean position.

A complete forecast consists of 12 h increments from 12-120 h (Figure 8). The forecast track valid at 0000 UTC 11 September 2010 consists of GE probability ellipses at each 12 h forecast interval. The GE ellipse is allowed to evolve with the dataset with no parameterizations or historical scaling to adjust the size, shape or orientation of the ellipse. The resulting ellipse displays the spatial variability in the forecast sequence that will be discussed in more detail in the Hurricane Earl case study.

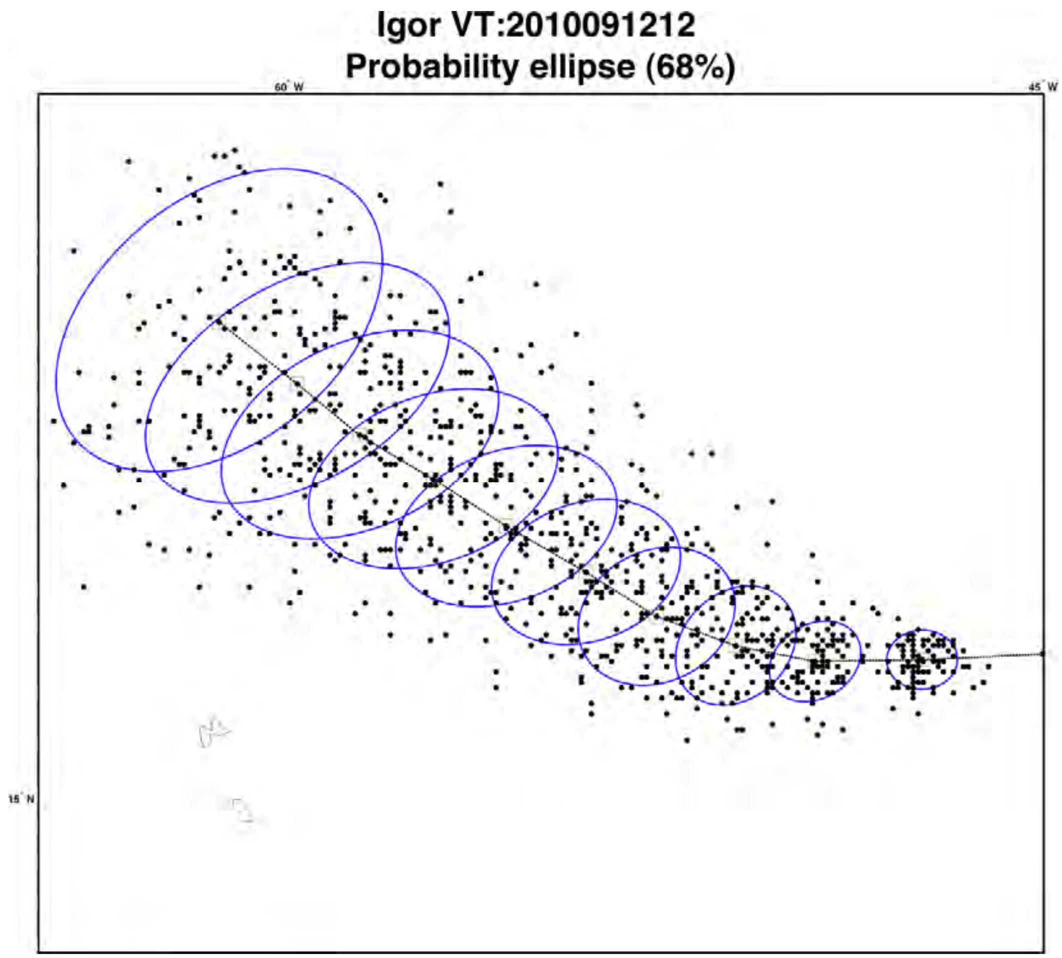

Figure 8. Ellipses defined to contain 68% of the GE members for 12-120 h forecast for Hurricane Igor initiated at 0000 UTC 11 September 2010. The red box indicates the GE mean location, the black dots represent individual ensemble member positions and the blue is the GE ellipse.

IV. ANALYSIS AND RESULTS

A. VALIDATING THE GE ELLIPSE

The performance of the GE ellipse is compared with the GPCE-X ellipses and the GPCE circles on the basis of reliability and sharpness. In this study, reliability is related to the hit rate associated with the proportion of time the best track TC position is inside the probability ellipse, and sharpness is defined as the size of the area contained by the ellipse.

1. Null Hypothesis

The null hypothesis to be tested is “The GE ellipse captures the best track verification position 68% of the time.” The alternative hypothesis is “The GE ellipse fails to capture the best track verification 68% of the time.” The null hypothesis is tested using the following procedure.

A simple test is conducted to determine if the best track position is located in the ellipse. If the ellipse contains (does not contain) the best track position, a hit (miss) is defined. The hits are then totaled and divided by the corresponding total of forecasts (sum of hits and misses) to determine the hit rate probabilities for each forecast interval. A Chi-squared test is then conducted to determine how closely the actual hit rate conforms to the predicted hit rate, which is defined by the 68% scale factor applied in construction of the ellipse. The distribution of hit rates for each forecast interval (Table 1) are found to not be statistically different than the predicted hit rate of 68%, therefore the null hypothesis “The GE ellipse captures the best track 68% of the time” is not rejected based on the results of the Chi-squared hypothesis test.

Table 1. 2008 through 2010 GE Hit percentages

Interval	12	24	36	48	60	72	84	96	108	120
GE hit %	80%	77%	73%	74%	74%	72%	68%	64%	65%	69%

2. Analysis of Spatial Variance

Neese (2010) identifies the need to divide the Atlantic basin into zones and consider probability density functions (PDF) for the TC center position in different regions of the basin. In his work, he considers the along-track and across-track variability of the PDF from different parts of the basin. He finds that the location should be considered for purposes of scaling probability cones. The spatial variability of the GE is considered qualitatively in this research based on the trajectory of the TC rather than the basin location.

Considered here is a qualitative discussion of the stochastic prediction of a TC embedded in a flow-dependent system, as related to the GE ellipse orientation. Hurricane Igor at the initial forecast time of 0000 UTC 12 September 2010 (Figure 9) consists of a westerly track over the first 60 h of the forecast period followed by slight recurvature after 60 h.

The latitudinal gradient of the steering flow can be inferred by considering the faster speed to the north and slower speed to the south of the TC forecast track while it was on a westerly trajectory. This is supported by the orientation of the ellipse. The flow north of the track is faster and the variance is shifted in a northwest to southeast orientation accordingly. The change in the orientation of the ellipses in Figure 9 from the 12 h to the 60 h interval indicates that the flow to the north of the ensemble mean is faster than the flow to the south of the ensemble mean. Therefore, in this scenario, a position slightly further east should be considered when a TC forecaster predicts the location of a TC on the south side of the ensemble mean. Likewise, if a TC location to the north of the ensemble mean is predicted, a position further west should be considered.

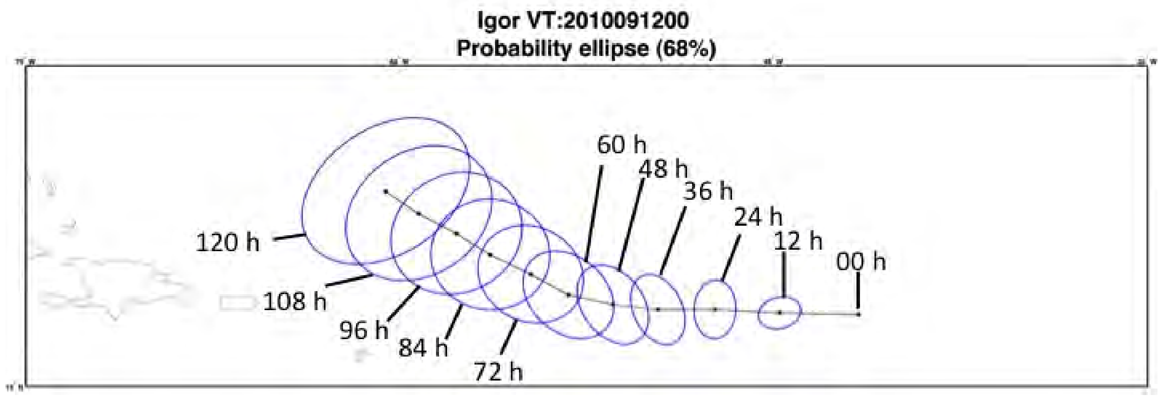

Figure 9. Hurricane Igor GE forecast for initial time 0000 UTC 12 September 2010

The recurvature of the TC is in response to changes in the flow pattern in which the TC is embedded. When the TC begins to recurve, the orientation of the ellipse changes and the areal extent increases indicating greater uncertainty in the forecast track. The predictions to the north of the ensemble mean recurve sooner than those to the south. Consistent with reduced speed during recurvature, the members to the right of the forecast mean slow and recurve while those to the left continued some time before they slow and recurv. By 84-hours the distribution is more normal to the mean trajectory, across-track error begins to dominate and the area of the GE probability ellipse increases. This is consistent with greater forecast track uncertainty during recurvature.

Generally, at times when the ellipse has an orientation other than normal to the ensemble mean track one can deduce that flow to one side of the ensemble mean forecast track is faster than the flow on the other side of the ensemble-mean forecast track. This change in orientation tells the more sophisticated user something about the flow in which the system is embedded and can be used to assist in selecting TC forecast positions in times when deviation from the GPCE circle is desired.

B. COMPARISON WITH EXISTING FORECAST TOOLS

1. Qualitative Comparison with GPCE

The GE probability ellipse has advantages and disadvantages when compared to GPCE circle. First, GPCE is debiased and parameterized based on historical averages.

These parameterizations ensure growth in the representation of the error with each forecast interval and ensure that the uncertainty area is consistent from forecast to forecast. Consistency is defined as ensuring the area of the probability circle doesn't grow and shrink dramatically between forecast intervals. Maintaining consistency in area is done to ensure forecaster confidence (Hansen et al. 2010).

As defined in the previous section, the anisotropic nature of the GE probability ellipse represents the spatial characteristics of the flow, which results in an ellipse oriented to the spatial variability of the individual ensemble members' predicted positions. The forecast track uncertainty is based on the different representations of the flow, which may reflect the possible origins of the uncertainty. This physical information is lost in the isotropic representation of the GPCE circle.

Further, because it is possible for predicted uncertainty to decrease with time, the GE probability ellipse is free to spatially vary in accordance with the evolution of the ensemble members. To demonstrate the feasibility of forecast uncertainty decreasing with time in a flow-dependent system, consider water flowing down a mountain during a rain event. The exact path a water parcel may follow from its point of impact with the earth is highly uncertain. However, predicting that the water will wind up in valley is not difficult. Another useful analogy is the weather ahead of and behind a cold front. The weather ahead of the system is unstable and to some degree uncertain. Behind the cold front the weather is highly predictable. This is an excellent example of increasing confidence in a prediction with time. One could consider Lorenz (1963) approach in which he identifies the atmosphere to be a quintessential chaotic system. There are certain points in the TC forecast track where it is in a highly stable part of the attractor and certain points when it is not. It is because of this that the GE ellipse was allowed to grow and occasionally shrink depending on the output of the EPS.

Particularly important are the changes in the along-track and across-track uncertainty. This is evident in the change of ellipse geometry throughout recurvature. While the change from a north-oriented to an east-oriented track occur for nearly all storms, the uncertainty associated with the timing and magnitude of this change remains large due to the variability in the mid-latitude flow pattern.

2. Forecast Track Error – Mean

The GPCE circle is centered on the consensus mean. Which is one of the important parameters in determining whether or not the GPCE circle captures the best track position. Similarly, the GE ellipse is centered on the mean of the ensemble members. Because both probability ellipses are centered on their respective means, it is important to determine how accurate the forecasts defined by each mean are relative to the best track (Figure 10). While it is clear from Figure 10 that the consensus mean is slightly more accurate than the GE mean the error bars indicating confidence intervals suggest that this difference was not statistically significant.

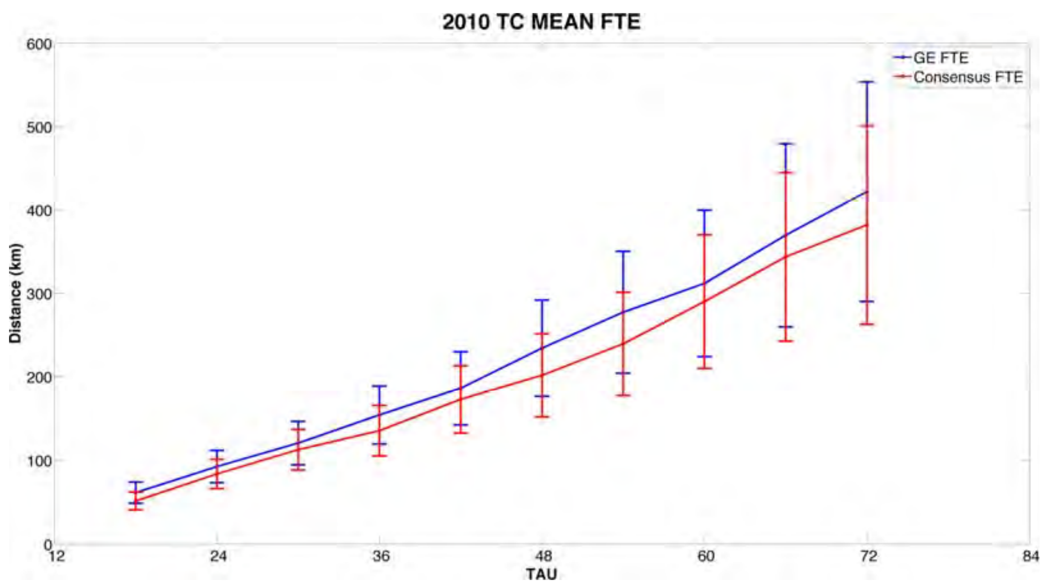

Figure 10. The mean forecast track error for the consensus mean (red) and the GE mean (blue). The error bars indicate confidence intervals of 95%.

Further analysis of the accuracy of the GE mean is conducted in which the forecasts based on the mean from each operational EPS means (e.g. ECMWF, MOGREPS, GFS) are considered. The results demonstrate that the accuracy of the individual EPS are not equal (Figure 11). Here it is evident that the ECMWF EPS mean is more accurate than the other model EPS means at forecast intervals past 72-hours. Whether this is a function of the higher resolution or the method of perturbing the statistics remains unclear, however the benefit to the GE mean becomes apparent later when relationship between hit-rates and sharpness is discussed.

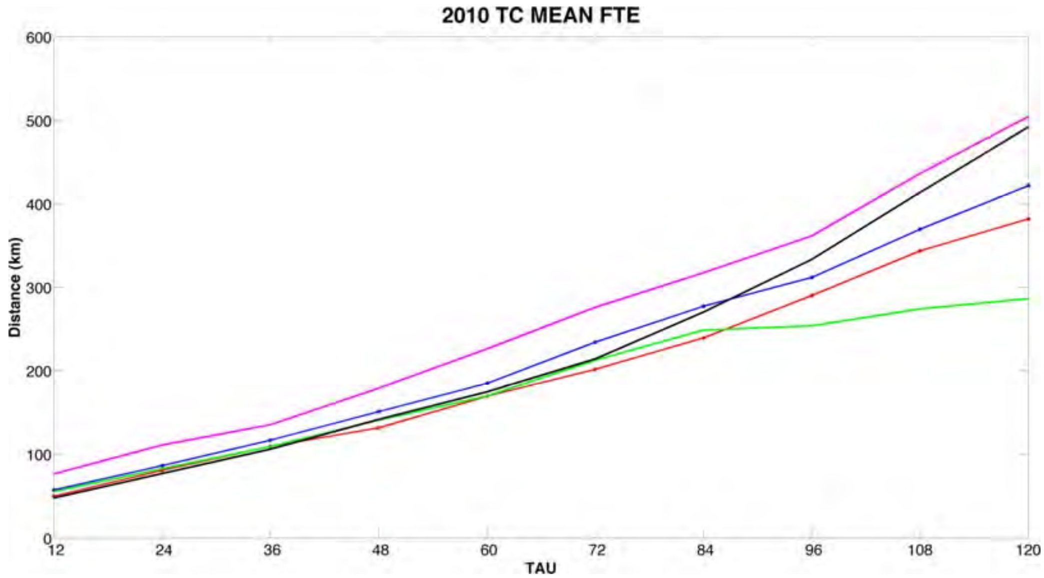

Figure 11. The EPS mean FTE from 12 h to 120 h forecast intervals. The MOGREPS mean is in magenta. The GFS mean is in black. The ECMWF mean is in green. The GE mean is in blue. The consensus mean is in red.

3. Sharpness – Mean Difference Area

Hansen et al. (2010) defines the mean area difference (MAD) between the GPCE and GPCE-X normalized by the GPCE area. The MAD is therefore a measure of comparative sharpness for the GPCE-X anisotropic ellipse specifically $\text{MAD} = (\text{GPCE}_{area} - \text{GPCE-X}_{area}) / \text{GPCE}_{area}$.

Following this methodology, the sharpness of the GE probability ellipse is evaluated using MAD for all homogenous forecasts for storms in 2010. The MAD is calculated to compare the sharpness of the GPCE probability circle and the GE probability ellipse. As stated previously, the years 2008 and 2009 are excluded due to significant changes in the GPCE format over those years. From 12 h to 36 h forecast intervals, the negative values for MAD (Table 2) are expected because of the perturbations applied in the analyses of the EPS models in the GE to produce unstable vectors in the ensembles. By the 36 h forecast interval, the GE sharpness matches that of the GPCE circle. For forecast intervals beyond 48 h, the GE produces a sharper forecast than the GPCE (Table 2).

Table 2. Mean Area Difference (Sharpness Comparison)

Forecast Interval	12	24	36	48	60	72	84	96	108	120
MAD(GPCE-GE) %	-66	-17	4	13	12	13	17	17	26	20

4. Reliability – Hit Rate

Reliability is defined as the match between the ellipse's confidence level (i.e. 68%) and the percentage of the time the best track is observed to fall within the ellipse. Hence, Majumbdar and Finocchio (2010) are correct–if the probability ellipse is perfectly reliable it would capture the best track 68% of the time. A forecast system in which the probability circle contains the best track more than 68% of the time is considered over-dispersive and under-dispersive if it is under 68%.

The GPCE circle hit rates are used as a basis of comparison for the GE probability ellipse (Table 3). In general, the GE is found to be more reliable than the GPCE (Table 3). The GE is over-dispersive in the first 24 h of the forecast, becomes more reliable through the 72 h forecast, and then slightly over-dispersive again out to 120 h. The over-dispersive tendency of the early forecast intervals is most likely due to the perturbed initial conditions of the EPS that made up the GE. However, the GE may benefit from the superior accuracy of the ECMWF EPS mean, which leads to more accurate forecast positions at later intervals and yields in the slightly over-dispersive results beyond 72 h.

Table 3. Percent of the time the 68% probability ellipses captures the best track position in 2010.

Tau	12	24	36	48	60	72	84	96	108	120
GPCE	71%	69%	73%	78%	81%	78%	80%	78%	77%	78%
GE	80%	78%	71%	65%	68%	66%	75%	72%	72%	77%

5. Summary of Statistics

To help determine what leads to a hit in the GPCE circle versus a hit in the GE ellipse, the MADs are plotted against the difference in FTE between the consensus mean and the GE mean (Figure 12). The red lines in Figure 12 divide the graph into quadrants at the zero axes.

In the upper right quadrant, there are instances when the GPCE circle captures the best track position and the GE ellipse does not. However, the opposite condition also occurs, in which the GE ellipse captures the best track location and the GPCE circle does not. In this quadrant, the GE ellipse is sharper than the GPCE circle, and the GE mean FTE is less than the consensus FTE. It is evident that the GPCE circle is not as sharp, because even though the consensus FTE is greater than the GE mean FTE the GPCE circle continues to capture the best track position. This is consistent with the over-dispersive tendency of the GPCE circle outlined in Table 3. The parameterization that forces the GPCE circle to grow with time may be responsible for this phenomenon. A more detailed analysis indicates that the when FTE is greater than +50 nm only hits from the GPCE intervals greater than 72 h are present.

In the upper left quadrant, the GPCE circle produces a sharper forecast than the GE ellipse, and the consensus FTE is smaller than the GE mean FTE. In lower left quadrant the GE ellipse is not as sharp as the GPCE circle and the GE mean FTE is greater than the consensus FTE. However, the GE ellipse sharpness does not decrease enough to compensate for the GE mean FTE. In lower right quadrant, where the consensus mean FTE is larger than the GE FTE, the GE sharpness decreases enough to capture the best track positions.

The performance of the GE compared to the GPCE in each of these quadrants indicates that FTE is a major factor in determining hits. Additionally, it is evident in the upper right quadrant that the parameter forcing the GPCE circles to consistently increase in areal extent may be the cause of the tendency to be over-dispersive as the hits for GPCE only beyond the FTE difference of +50 nm are from forecast intervals beyond 72 h.

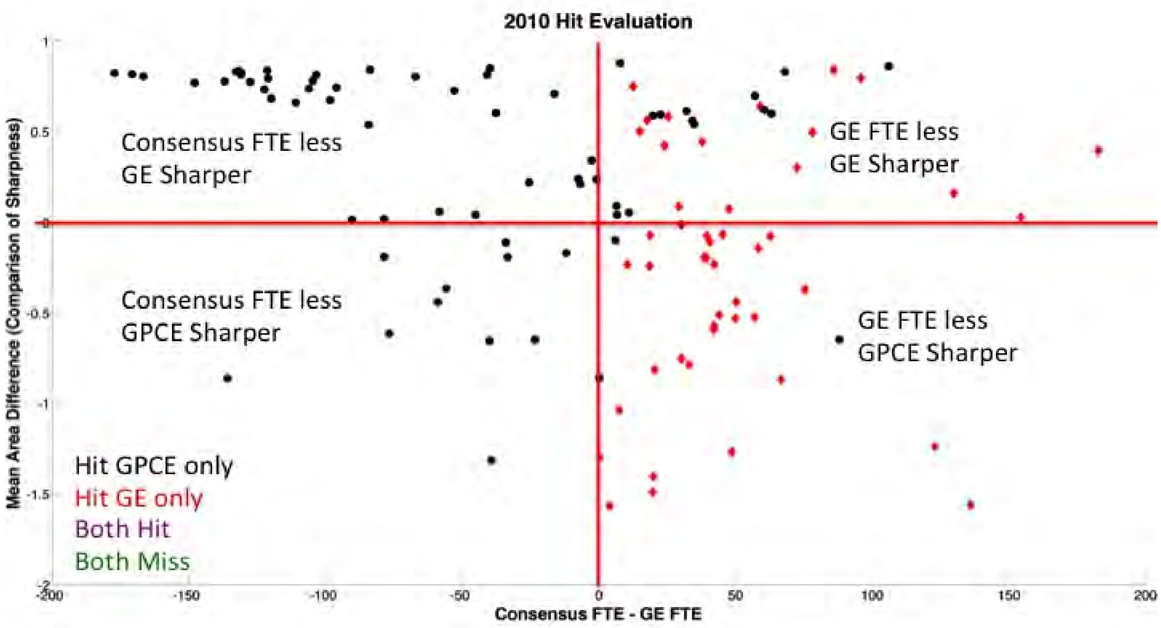

Figure 12. Mean Area Difference versus FTE difference. Red indicates GE only hits. Black indicates GE only hits. Green indicates no hits. Magenta indicates both hits.

6. Additional Statistics (GE, GPCE and GPCE-X comparison)

While the primary focus of this thesis is to compare the GE probability ellipse to the GPCE circle, the GPCE-X statistics for 2010 are also available in an experimental mode. The comparisons of the GPCE-X and GE in Table 4 and Table 5 are calculated following the same procedures used for the GE and GPCE comparisons. The values from these comparisons are included for convenience.

The benefits of the anisotropic nature of the GPCE-X ellipse are evident in the improvement of the sharpness (Table 4) and reliability (Table 5) of the GPCE-X ellipse over the GPCE circle. Therefore, the GPCE-X ellipse and GE ellipse are both sharper than the GPCE circle.

The primary differences in the sharpness between the GPCE-X and the GE occur at the short and long-range forecast intervals. For 12 h and 24 h forecasts, the GE is not as sharp as GPCE-X and at 108 h and 120 h forecasts the GPCE-X is not as sharp as the GE (Table 4). The difference in the early forecast period is most likely due to the perturbations in the initial conditions used to define the ensemble. The differences later in the forecast period are most likely due to the scaled growth of the GPCE-X to ensure the growth of predicted forecast error with time.

The hit percentages for 2010 define the reliability of the probabilistic ellipse (Table 5). The GPCE is persistently over-dispersive for all forecast intervals, and the GPCE-X is reliable through all forecast intervals. As discussed earlier the GE is over dispersive at the short forecast intervals and at 120 h.

Table 4. Mean Area Difference 2010

Forecast Interval	12	24	36	48	60	72	84	96	108	120
MAD(%)= $\frac{(\text{GPCE - GPCE-X})}{\text{GPCE}}$	-18	-11	0	13	13	11	13	13	14	16
MAD(%)= $\frac{(\text{GPCE - GE})}{\text{GE}}$	-66	-17	4	13	12	13	17	17	26	20
MAD(%)= $\frac{(\text{GPCE-X - GE})}{\text{GPC-X}}$	-51	-9	1	0	-4	-2	2	2	13	6

Table 5. Hit percentages 2010

Forecast interval	12	24	36	48	60	72	84	96	108	120
GPCE (68%)	71	69	73	78	81	78	80	78	77	78
GPCE-X (70%)	69	64	71	67	77	64	72	63	65	67
GE (68%)	80	78	71	65	68	66	75	72	72	77

Although the GPCE-X and the GE statistics are essentially equal, it is important to remember the key difference between the two methods. The GPCE-X is based on the consensus of deterministic models. Because of this it can have as few as only two models at the time of rendering. By convention, it is oriented with the semi-major axis either normal to or parallel to the trajectory. This is a significant step forward from the isotropic circle, but it may not capture the flow-dependent nature of the uncertainty. The GE does convey information to the user about the physics of the flow-dependent system in which the TC is embedded.

THIS PAGE INTENTIONALLY LEFT BLANK

V. CASE STUDY EARL

A. INTRODUCTION

Hurricane Earl was selected as a case study because of the interest in the successful decision not to sortie the Atlantic Fleet from Norfolk VA as Earl approached the area. As stated in the introduction, the Second Fleet Oceanographer made the recommendation to not sortie the fleet based on personal confidence in the interpretation of the NHC forecasts, which he augmented by the use of probability products that supported the recommendation and helped convey forecast confidence levels. This decision to not sortie has been used as an example of why probability products are important in communicating forecast confidence to commanders.

On 25 August 2010, TD 07 formed in the eastern Atlantic. The depression intensified into TS Earl by 1200 UTC 25 August 2010. Earl followed a westerly track and intensified only slightly until the storm intensified into a hurricane. On 29 August 2010, Earl underwent rapid intensification and reached category four strength as it approached the U.S. Virgin Islands and Puerto Rico. Hurricane Earl tracked northwest passing just north of the Leeward Islands, and began a period of gradual weakening. The hurricane recurved east of North America, and made its closest approach to the U.S. mainland as it skirted approximately 140 km seaward of Cape Hatteras.

B. FORECAST EVOLUTION OF HURRICANE EARL

1. Hurricane Earl 1200 UTC 28 August 2010

The Hurricane Earl case study began with 1200 UTC 28 August 2010 (Figure 13). Early in the forecast period, the GPCE circle was clearly sharper. The GE ellipse was most likely larger at these early forecast intervals due to the perturbations used in the ensembles. However, by 48 h the GPCE circles covered more area than the GE ellipses indicating that the GE ellipse is now producing a sharper forecast.

The GE ellipses also provided information about the predicted spatial uncertainty associated with the flow. Early in the forecast period the GE members to the left of the mean were moving faster than those to the right. This was most likely because the members on the right of the mean have slowed and began recurvature sooner than those to the left of the mean. By 84 h, the trajectory of Earl passed through the heading of 45 degrees. Those members that recurved early were controlled by faster flow to the right of the storm, and they became embedded in the stronger mid-latitude westerly flow. In response, the ellipse elongated along-track. Additionally, by 120 hours the GPCE circle was clearly covering a much larger area than the GE ellipse. For a forecaster, the orientation of the uncertainty communicated confidence in the forecast both along and across-track.

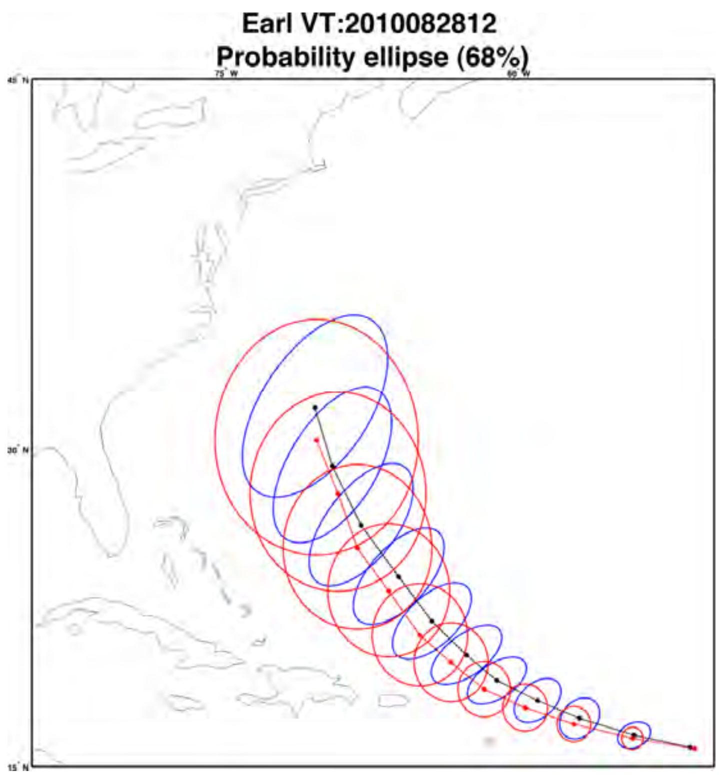

Figure 13. The forecast for Hurricane Earl initiated at 1200 UTC 28 August 2010 for 0 h through 120 h in 12-h increments. The red circles define the GPCE, and the GE probability ellipse is shown in blue.

2. Hurricane Earl 0000 UTC 30 August 2010

The forecast track for Earl initiated at 0000 UTC 30 August 2010 (Figure 14) exhibited a significant along-track uncertainty. This was indicative of an uncertainty in the speed of Earl and was an important consideration when determining the last chance to sortie the fleet. In situations like this, the preference for the ellipse over a swath type product was clear. A swath would not adequately communicate the along-track uncertainty. Similarly, the advantage of the anisotropic nature of the GE ellipses was obvious. The large along-track uncertainty would have resulted in a false representation of across-track uncertainty if the uncertainty were summarized isotropically.

A more detailed look at the individual model performance illustrated the reason for the large spatial coverage of the GE ellipse at this forecast time. The MOGREPS model was much slower in recurvature than the GFS and the ECMWF. This was primarily a result MOGREPS building the subtropical high farther into the southwest Atlantic resulting in a late recurvature.

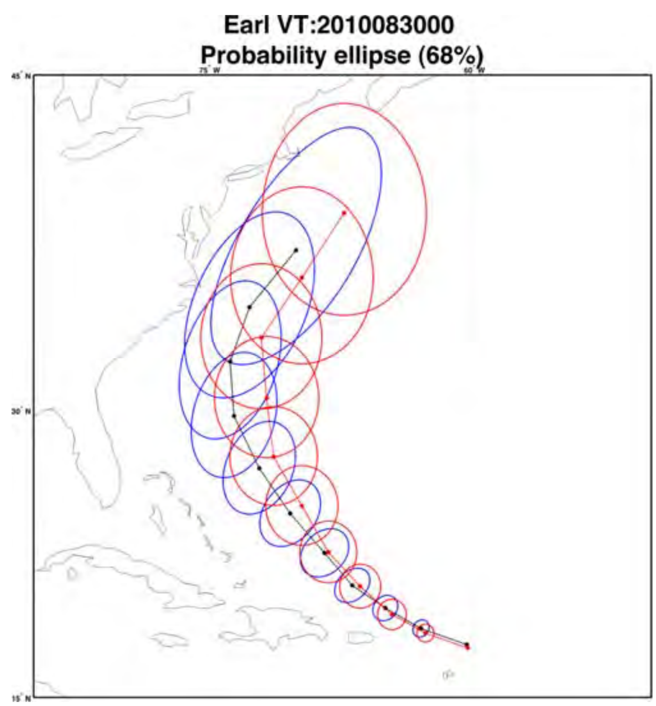

Figure 14. The forecast for Hurricane Earl initiated at 0000 UTC 30 August 2010 for 0 h through 120 h in 12-h increments. The red circles define the GPCE, and the GE probability ellipse is shown in blue.

3. Hurricane Earl 0000 UTC 31 August 2010

By 0000 UTC 31 August 2010 (Figure 15), the TC track was forecast to pass Norfolk, VA within 84 hours. The GPCE forecast was slightly sharper than the GE, however, the GE conveyed that the uncertainty was clearly along-track. Norfolk VA was inside, but on the outer edge of the 68% probability circle of GPCE, and just outside of the GE probability ellipse. This forecast was issued 84 hours before Hurricane Earl would make its closest approach to Norfolk affording the operational commanders another 24 hours before making the final sortie decision. Regnier and Harr (2006) demonstrated that waiting until the most current information was available for making the final sortie recommendation was desirable. However, the leading edge of the probability ellipse at 72-hours should have been considered as a worst-case scenario.

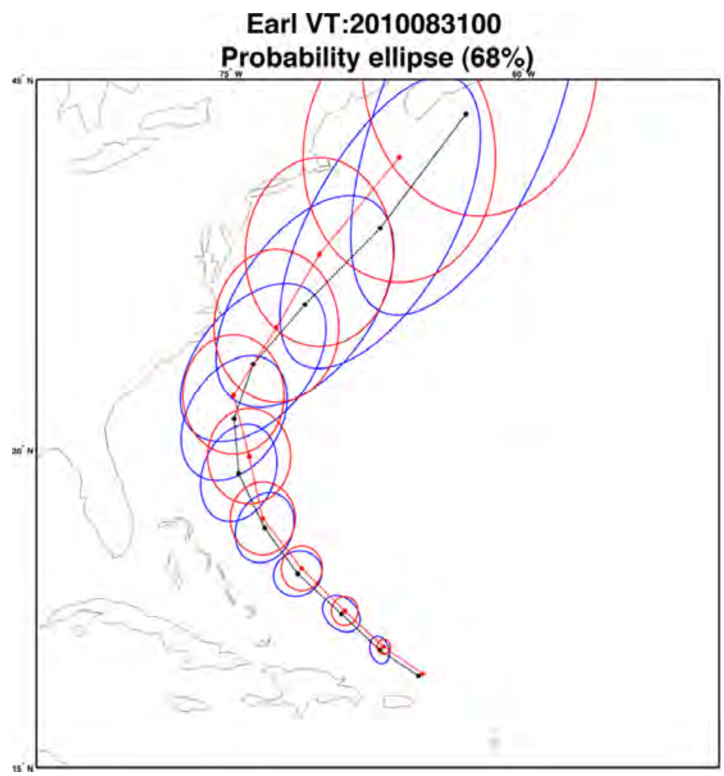

Figure 15. The forecast for Hurricane Earl initiated at 0000 UTC 31 August 2010 for 0 h through 120 h in 12-h increments. The red circles define the GPCE, and the GE probability is shown in blue.

4. Hurricane Earl 1200 UTC 31 August 2010

Sortie condition Bravo, which requires ships to make preparations to sortie within 24 hours, was set in conjunction with the forecast at 1200 UTC 31 August 2010 (Figure 16). This warning gave advance notice to the ship Captains to have their Sailors make preparations to get underway while still allowing time to ensure safety for their families. At this forecast time the 72 h GPCE circle contained Norfolk, VA Navy Base. However, the GE continued to showed along-track uncertainty as the dominant predicted error in Hurricane Earl's track.

At this stage of the forecast track, the isotropic nature of the GPCE circle masked the uncertainty in the flow pattern. Though it was possible that the uncertainty in the models that the make up the GPCE consensus was along-track, the GPCE circle could not depict this condition. The spatial variability of the GE ellipse clearly indicated the semi-minor axis in the across-track direction. Thus, indicated a higher degree of confidence in the storm staying seaward of Norfolk.

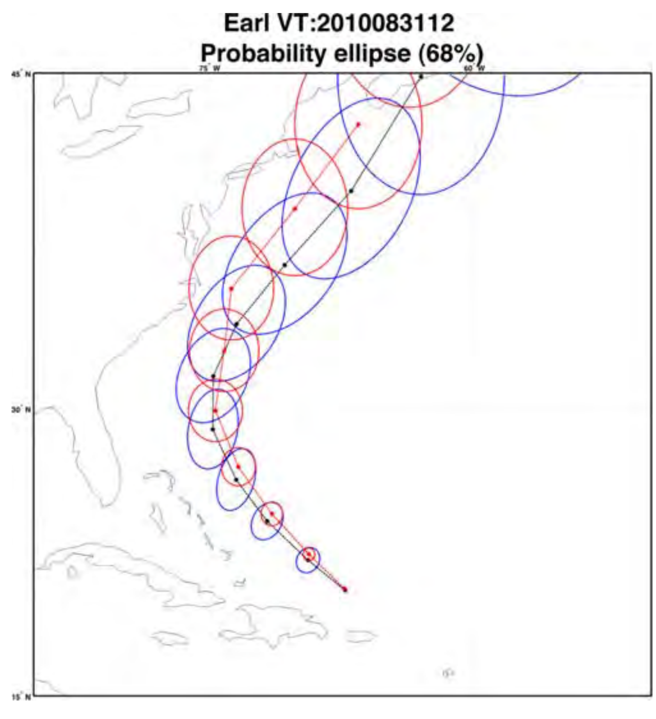

Figure 16. The forecast for Hurricane Earl initiated at 1200 UTC 31 August 2010 for 0 h through 120 h in 12-h increments. The red circles define the GPCE, and the GE probability is shown in blue.

5. Hurricane Earl 0000 UTC 01 September 2010

By 0000 UTC 01 September 2010 (Figure 17) it was evident from the area of the GE probability ellipses and the GPCE circles that Earl would stay offshore. The decision was made to not sortie the fleet. The GE ellipse supported this decision and indicated that in this scenario the tool could have been used to help make the decision in two key ways. First, the uncertainty was decidedly along-track, so confidence in the storm missing Norfolk was high. Second, if it was deemed necessary to sortie the leading edge of the GE ellipse served as a good worst-case scenario for storm location. Further analysis of a larger set of storms in a detailed decision analysis framework is necessary to make more definitive statements about the usefulness of these tools in a sortie context.

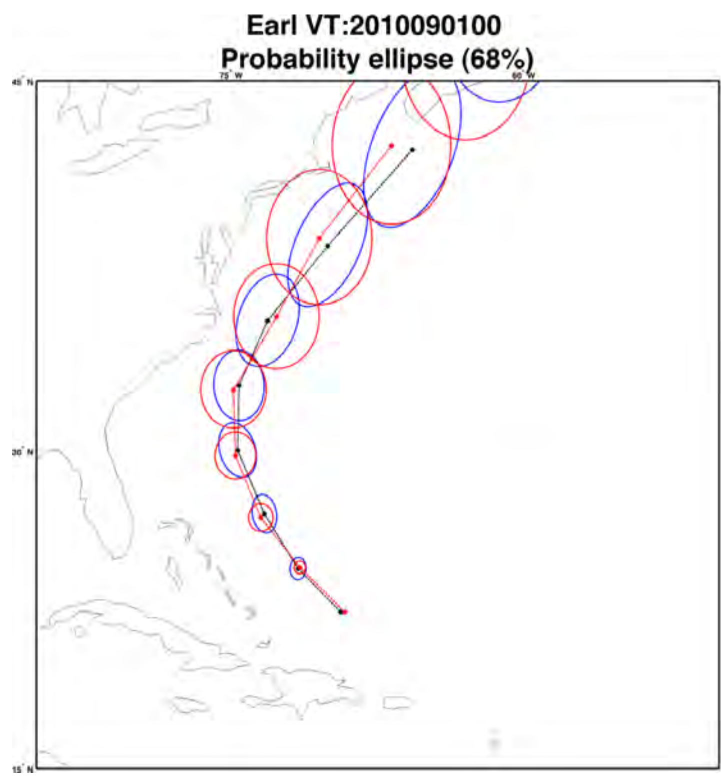

Figure 17. The forecast for Hurricane Earl initiated at 1200 UTC 28 August 2010 for 0 h through 120 h in 12-h increments. The red circles define the GPCE, and the GE probability is shown in blue.

V. CONCLUSIONS AND RECOMMENDATIONS

A. UTILITY OF A GE ELLIPSE

This thesis demonstrates that the GE probability ellipse is a good predictor of forecast uncertainty. The GE ellipse is a useful addition to process of indicating uncertainty in TC track forecasts, because the GE ellipse retains information about the physics of the flow through the spatial variability of the ellipse along the forecast track. The qualitative assessment of the spatial variability shows that the GE ellipse can be used to support decisions that depend on TC speed forecast when left or right of the ensemble mean. It also helps determine the degree of uncertainty in the area of recurvature.

The GE probability ellipse reliability tests indicate that the 68% GE ellipse is reliable, i.e. it containing the best track position near 68% of the time for the 2008 through 2010 seasons. The GE probability ellipse forecast sharpness was also shown to be superior to the GPCE circles at times beyond the 36-hour forecast interval. Thus, the GE provides end users with a sharp and reliable tool to augment existing probabilistic forecasting tools.

The anisotropic nature of the GE ellipse makes it particularly useful. The ability to distinguish between along and across-track uncertainty assists a forecaster in communicating confidence in a forecast track. Additionally the spatial variability of the GE ellipse demonstrates the retention of some information about the physics of the flow-dependent system in which the TC is embedded. This communicates information to the sophisticated user. The examples of the gradient of flow across the track from Hurricane Earl and Hurricane Igor demonstrate potential growth in the utility of this tool.

Two potential audiences or users for the GE ellipse exist. Each possesses a different level of sophistication. Operational TC forecasting centers could use the ellipse for TC track prediction and the Fleet METOC officers could use the GE ellipse to understand uncertainty in TC forecasts when making sortie recommendations.

Operational TC forecast centers currently use the GPCE circle as the best first choice for TC forecast track position placement. The GE ellipse can assist TC forecasters in understanding instances when deviation from the GPCE should be made. Since a measure of the physics of the flow-dependent system is retained in the behavior of the GE ellipse it can serve as a robust support tool for persuading one to deviate or not deviate from the GPCE constraints.

The GE ellipse by definition is a prediction of forecast uncertainty representing one standard deviation of the combined ensembles' representation of the future atmosphere. Consequently, it can be used by a Fleet Oceanographer to support sortie recommendations. The GE ellipse provides a quantitative tool to answer the question of "how confident are you in this forecast" that inevitably comes with the approach of a TC. Fleet Oceanographers are considered sophisticated users for purposes of this study. It is assumed they have a background in Meteorology and Oceanography, and that they possess the basic skills to interpret the GE ellipse. Under this assumption it is reasonable to expect the Fleet METOC Officer to use such guidance to assist in making recommendations. Influencing sortie recommendations is at the top of the Battle Space on Demand (BoND) pyramid and is the ultimate goal of the Meteorology and Oceanography program sponsor– and it is the ultimate goal of forecasting. To reduce uncertainty, or at least quantify the uncertainty surrounding operational decisions is the goal of operational METOC officers. The GE ellipse can assist in this capacity.

B. RECOMMENDATIONS

In the case study the probability product used by the Second Fleet Oceanographer to support the recommendation to not sortie displayed the probability of destructive (50 kt) winds. The possible positions of the TC used in this product are found by sampling from a PDF of the five-year historical error contained in the NHC cone of uncertainty scaled based on the size of the GPCE circle. This cone of uncertainty is static and isotropic about the official forecast track. It may be that the distribution of the GE ellipse provides a better near real-time PDF from which to sample.

The positive results of the qualitative analysis of the spatial variability of the ellipse indicates further need for a quantitative analysis of performance based on the trajectory of the storm and location with the basin. Detailed analysis in theses areas will help define the utility of the GE ellipse as a tool at both levels and identify locations where the distribution could be sampled from to develop more robust wind speed and storm surge probability products.

A qualitative assessment of the performance of the GE mean and spread demonstrates that the spread of the ensembles contains the best track through most cases. In 2010, with the exception of Hurricane Earl, the forecast tracks were captured well within the spread. The speed was the largest source of uncertainty since small across-track errors can result in large along-track errors depending on the gradient of the flow velocity. Future investigation into bias corrections at various places in either space of storms track could correct for this speed error and further improve on the skill of the GE mean.

Study of the variability of the area compared to the trajectory of the storm in a similar fashion to the basin ATE and XTE research could yield additional insight into the evolution of uncertainty through the TC life cycle. This may also result in a more precise PDF from which to sample when producing probability products.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Bougeault, Phillippe, and Coauthors, 2010: The Thorpex interactive grand global ensemble. *Bull. Am. Meteor. Soc.,* **91**,1059-1072.

Bowler N. E, A. Arribas, K. R. Mylne, K. B. Robertson and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.,* **134**,703–722.

Carr, L. E. and R. L. Elsberry R. L., 2000: Consensus of dynamical tropical cyclone track forecasts-error versus spread. *Mon. Wea. Rev.* **128**, 4131–4138.

DeMaria, M., J. A. Knaff , R. Knabb, C. Lauer, C. R. Sampson, R. T. DeMaria, 2009: A New Method for Estimating Tropical Cyclone Wind Speed Probabilities. *Wea. Forecasting,* **24**,1573-1591.

Dorsey, J., 2003: Navy Costs for Isabel At Least $105.6 Million. 27 September 2009, *Norfolk Virginian-Pilot*.

ECMWF, cited 2011: Predictability: The ECMWF Ensemble Prediction System. [Available online at http://www.ecmwf.int/research/predictability/projects/index.html] accessed 12 June 2011

Goerss, J S. 2000: Tropical cyclone track forecasts using an ensemble of dynamic models. *Mon. Wea. Rev.*, **128**, 1187–1193.

Goerss, J S. 2007: Prediction of Consensus Tropical Cyclone Track Forecast Error. *Mon. Wea. Rev.*, **135**, 1985–1993.

Hansen, J. A., J. S. Goerss, C. Sampson, 2010: GPCE-AX: An Anisotropic Extension to the Goerss Consensus Error. *Wea. Forecasting*, **26**, 416-422.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **42**, 433–471.

Majumbar, S. J. and P. M. Finocchio, 2009: On the ability of global ensemble prediction systems to predict tropical cyclone track probabilities *Wea. Forecasting*, **25**, 659–680.

Neese, J. M., 2010: Evaluating Atlantic tropical cyclone track error distributions for use in probabilistic forecast of wind distribution. M. S. Thesis, Naval Postgraduate School, 65 pp. [Available from http://edocs.nps.edu/].

NHC, cited 2010a: Tropical Weather Summary. [Available online at http://www.nhc.noaa.gov/archive/2007/tws/MIATWSAT_nov.shtml] accessed 14 December 2010.

NHC, cited 2010b: Uncertainty Cone. [Available online at http://www.nhc.noaa.gov/aboutcone.shtml] accessed 10 July 2011.

Regnier, E. and P. A. Harr, 2006: A Dynamic Decision Model Applied to Hurricane Land Fall. *Wea. Forecasting*, **21**, 764–780.

Sampson, C. R., J. A. Knaff, and E. M. Fukada, 2006: Operational Evaluation of a Selective Consensus in the Western North Pacific Basin. *Wea. Forecasting*, **22**, 671–675.

UCAR, cited 2010: NCEP Perturbation Method. [Available online at http://www.meted.ucar.edu/nwp/pcu2/ens_matrix/mref_pert1.htm] accessed 14 May 2011.

U.S. Navy, cited 2010a: Navy facilities within the U.S.[Available online at http://www.navy.mil/navydata/bases/navbases.html] accessed on 15 December 2010.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. William P. Roeden
   45th Weather Squadron Meteorologist
   Systems Staff Meteorology
   45 WS/SYS
   1201 Edward H. White II St.
   MS 7302
   Patrick AFB, Florida 32925-3238

4. Director Joint Typhoon Warning Center
   JTWC
   Pearl Harbor, Hawaii

5. Richard Jeffries
   Commander Naval Meteorology and Oceanography Command
   N4/7
   Stennis Space Center, Massachusetts

6. Superintendent
   Naval Research Laboratory
   Monterey, California

7. Dr. Russ Elsberry
   Naval Postgraduate School
   Meteorology Department
   Monterey, California